A Generic Method for the Production of Recombinant Proteins in *Escherichia coli* Using a Dual His_6-MBP Affinity Tag

Joseph E. Tropea,[1] Scott Cherry,[1] Sreedevi Nallamsetty,[1]

Christophe Bignon[2] and David S. Waugh,[1,*]

[1] Macromolecular Crystallography Laboratory, National Cancer Institute at Frederick, P.O. Box B, Frederick, MD 21702-1201. [2] AFMB, UMR 6098, CNRS & Universites Aix-Marseille I & II, 31 Chemin J. Aiguier, 13402 Marseille Cedex 20, France.

*Corresponding author. Tel: (301) 846-1842; Fax: (301) 846-7148; Email: waughd@ncifcrf.gov

1. Introduction

Because of its remarkable ability to enhance the solubility and promote the proper folding of its fusion partners, *Escherichia coli* maltose-binding protein (MBP) has emerged as an attractive vehicle for the production of recombinant proteins (***1,2***). However, MBP fusion proteins do not always bind efficiently to amylose resin, and even when they do amylose affinity chromatography typically does not produce samples of sufficient purity for structural studies. To address this problem, we identified several locations in which small affinity tags can be inserted within the framework of an MBP fusion protein without compromising its solubility-enhancing properties (***3***). Here, we describe how one such configuration, in which a hexahistidine tag (His_6) is added to the N-terminus of MBP, forms the foundation of an entirely generic strategy for protein production in *E. coli*. The MBP moiety improves the yield and enhances the solubility of the passenger protein while the His-tag facilitates its purification. The soluble fusion protein (His_6-MBP-passenger) is purified by immobilized metal affinity chromatography (IMAC) on Ni-NTA resin and then cleaved in vitro with His_6-tagged tobacco etch virus protease (His_6-TEV protease) to separate the His_6-MBP from the passenger protein. In the final step, the unwanted byproducts of the digest, as well as any impurities that eluted from the Ni-NTA resin along with the fusion protein in the first IMAC step, are absorbed by a second round of IMAC, leaving nothing but the pure passenger protein in the flow-through fraction. Hence, the application of two successive IMAC steps, rather than just one, is the key to obtaining crystallization-grade protein with a single affinity technique. This simple generic protocol should be readily amenable to automation for high-throughput applications.

2. Materials

2.1. Recombinational Vector Construction

1. The Gateway® destination vector pDEST-HisMBP (*see* **Fig. 1**).
2. Reagents and thermostable DNA polymerase for PCR amplification (*see* **Note 1**).
3. Synthetic oligodeoxyribonucleotide primers for PCR amplification (*see* **Fig. 2**).
4. TE buffer (10 mM Tris-HCl [pH 8.0], 1 mM EDTA)
5. TAE-agarose and an apparatus for submarine gel electrophoresis of DNA (*see* **Note 2**).
6. QIAquick™ gel extraction kit (Qiagen, Valencia, CA, USA) for the extraction of DNA from agarose gels.
7. Chemically competent DB3.1 cells (Invitrogen, Carlsbad, CA, USA) for propagating pDEST-HisMBP and pDONR201.
8. Competent *gyrA*$^{+}$ cells (*e.g.,* DH5□, MC1061, HB101) (*see* **Note 3**)
9. Gateway® PCR Cloning System (Invitrogen, Carlsbad, CA, USA).
10. LB medium and LB agar plates containing ampicillin (100 □g/ml). LB medium: Add 10 g bacto tryptone, 5 g bacto yeast extract, and 5 g NaCl to 1 liter of H_2O and sterilize by autoclaving. For LB agar, also add 12 g of bactoagar before autoclaving. To prepare plates, allow medium to cool until flask or bottle can be held in hands without burning, then add 1 ml ampicillin stock solution (100 mg/ml in H_2O, filter sterilized), mix by gentle swirling, and pour or pipet ca. 30 ml into each sterile petri dish (100 mm dia.).
11. Reagents for small-scale plasmid DNA isolation (*see* **Note 4**).

12. An incubator set at 37 ºC.

2.2 Pilot Expression Experiment

1. Competent BL21PRO cells (Clontech, Palo Alto, CA, USA) containing the TEV protease expression vector pRK603 (***4***) (*see* **Notes 5** and **6**).
2. A derivative of pDEST-HisMBP that produces a His_6-MBP fusion protein with a TEV protease recognition site in the linker between MBP and the passenger protein (*see* **Subheading 3.1.**).
3. LB agar plates and broth containing both ampicillin (100 μg/ml) and kanamycin (25 μg/ml). See **Subheading 2.1.10.** for LB broth, LB agar, and ampicillin stock solution recipes. Prepare stock solution of 25 mg/ml kanamycin in H_2O and filter sterilize. Store at 4°C for up to 1 month. Dilute antibiotics 1000-fold into LB medium or molten LB agar.
4. Isopropyl-thio-β-D-galactopyranoside (IPTG), analytical grade (Anatrace Inc., Maumee, OH, USA). Prepare a stock solution of 200 mM in H_2O and filter sterilize. Store at -20 °C.
5. Anhydrotetracycline (ACROS Organics/Fisher Scientific, Springfield, NJ, USA). Prepare a 1000X stock solution by dissolving in 50% ethanol at 100 μg /ml. Store in a foil-covered tube at –20 ºC.
6. Shaker/incubator.
7. Sterile baffle-bottom flasks (Bellco Glass, Inc., Vineland, NJ, USA)
8. Cell lysis buffer: 20 mM Tris-HCl (pH 8), 1 mM EDTA.
9. Sonicator (with microtip).

10. 2x SDS-PAGE sample buffer (Invitrogen, Carlsbad, CA, USA) and 2-mercaptoethanol (Sigma Chemical Co., St. Louis, MO, USA).
11. SDS-PAGE gel, electrophoresis apparatus, and running buffer (*see* **Note 7**).
12. Gel stain (*e.g*. Gelcode® Blue from Pierce, Rockford, IL, USA, or PhastGel™ Blue R from Amersham Biosciences, Piscataway, NJ, USA).
13. Spectrophotometer
14. 1.5 ml microcentrifuge tubes

2.3. Large-Scale Cell Growth and Protein Purification

2.3.1. Cell Growth.

1. LB broth (*see* **Subheading 2.1.10**).
2. Sterile 500 ml and 4 liter baffled-bottom flasks (Bellco Glass, Inc., Vineland, NJ, USA).
3. Sterile ampicillin solution (100 mg/ml, *see* **Subheading 2.1.10**).
4. Competent BL21 cells (Novagen, Madison, WI, USA).
5. D(+)-Glucose monohydrate, BioChemika grade (Fluka Chemical Corp., Milwaukee, WI, USA).
6. IPTG (*see* **Subheading 2.2.4**).
7. A derivative of pDEST-HisMBP that produces a His_6-MBP fusion protein with an intervening TEV protease recognition site.
8. Shaker/incubator.
9. Spectrophotometer.

2.3.2. Protein Purification.

1. ÄKTA Explorer chromatography system (Amersham Biosciences, Piscataway, NJ, USA).
2. Ni-NTA Superflow (Qiagen Inc, Valencia, CA, USA).
3. Column XK 16/10 (Amersham Biosciences, Piscataway, NJ, USA).
4. 0.22 μm Polyethersulfone filter (Corning Inc., Corning, NY, USA).
5. HEPES, BioChemika grade (Fluka Chemical Corp., Milwaukee, WI, USA).
6. Imidazole, BioChemika grade (Fluka Chemical Corp., Milwaukee, WI, USA).
7. Sodium chloride, ACS reagent grade (J. T. Baker, Phillipsburg, NJ, USA).
8. 10 *N* sodium hydroxide solution, reagent grade (J. T. Baker, Phillipsburg, NJ, USA).
9. Hydrochloric acid, ACS plus grade (Fisher Scientific, Pittsburgh, PA, USA).
10. Benzamidine hydrochloride: hydrate (Sigma Chemical Co., St. Louis, MO, USA).
11. Complete EDTA-free protease inhibitor cocktail tablets (Roche Diagnostics GmbH, Mannheim, Germany).
12. SDS-PAGE gel, 2X SDS-PAGE sample buffer, electrophoresis apparatus, and running buffer (*see* **Subheadings 2.2.10., 2.2.11. and 2.2.12.**).
13. His_6-TEV protease (***5***).

3. Methods

3.1. Construction of His_6-MBP Fusion Vectors by Recombinational Cloning

The Gateway® recombinational cloning system is based on the site specific recombination reactions that mediate the integration and excision of bacteriophage lambda into and from the *E. coli* chromosome, respectively. For detailed information about this system, the investigator is encouraged to consult the technical literature supplied by Invitrogen, Inc.

3.1.1. pDEST-HisMBP

To utilize the Gateway® system for the production of His_6-MBP fusion proteins, one must first construct or obtain a suitable "destination vector". Currently there are no commercial sources for such vectors. An example of a destination vector that can be used to produce His_6-MBP fusion proteins (pDEST-HisMBP), which is available from the authors, is shown in **Fig. 1**. pDEST-HisMBP was constructed by inserting an in-frame hexahistidine sequence between codons 3 and 4 of MBP in pKM596 (***6***).

The Gateway® cloning cassette in pDEST-HisMBP carries a gene encoding the DNA gyrase poison CcdB, which provides a negative selection against the destination vector, the donor vector, and various recombination intermediates so that only the desired recombinant is obtained when the end products of the recombinational cloning reaction are transformed into *E. coli* and grown in the presence of ampicillin. pDEST-HisMBP and other vectors that carry the *ccdB* gene must be propagated in a host strain with a *gyrA* mutation (*e.g. E. coli* DB3.1) that renders the cells immune to the action of CcdB.

3.1.2. Gateway® Cloning Protocol

The easiest way to construct a His_6-MBP fusion vector by recombinational cloning is to start with a PCR amplicon wherein the open reading frame (ORF) of interest is bracketed by *att*B1 and *att*B2 recombination sites on its N- and C-termini, respectively, which can be generated by amplifying the target ORF with PCR primers that include the appropriate *att*B sites as 5' unpaired extensions (*see* **Fig. 2**). The 3' ends of the PCR primers are chosen so that they will be able to form 20-25 base pairs with the template DNA. A recognition site for TEV protease is incorporated between the N-terminus of the ORF and the *att*B1 site in this PCR amplicon, enabling the passenger protein to be separated from the N-terminal His_6-MBP tag. Although this can be accomplished with a single, long N-terminal PCR primer for each gene, we normally perform the PCR amplification with two overlapping N-terminal primers instead, as outlined in **Fig. 2**. Two gene-specific primers (N1 and C) are required for each ORF. The C-terminal primer (C) includes the *att*B2 recombination site as a 5' extension. The 5' extension of the N-terminal primer (N1) includes a recognition site for TEV protease. The PCR product generated by these two primers is subsequently amplified by primers N2 and C to yield the final product. Primer N2 anneals to the TEV protease recognition site and includes the *att*B1 recombination site as a 5' extension. This generic PCR primer can be used to add the *att*B1 site to any amplicon that already contains the TEV protease recognition site at its N-terminal end. The PCR reaction is performed in a single step by adding all three primers to the reaction at once (*see* **Note 8**). To favor the accumulation of the desired product, the *att*B-containing primers are used at typical concentrations for PCR but the concentration of the gene-specific N-terminal primer (N1) is 20-fold lower.

1. The PCR reaction mix is prepared as follows (*see* **Note 9**): 1 μl template DNA (~10 ng/μl), 10 μl thermostable DNA polymerase 10X reaction buffer, 16 μl dNTP solution (1.25 mM each), 2.5 μl primer N1 (~1μM, or 13 ng/μl for a 40mer), 2.5 μl primer N2 (~20μM, or 260 ng/μl for a 40mer), 2.5 μl primer C (~20μM, or 260 ng/μl for a 40mer), 1 μl thermostable DNA polymerase, 64.5 μl H_2O (to 100 μl total volume).
2. The reaction is placed in the PCR thermal cycler with the following program: initial melt for 5 min at 94 ºC; 30 cycles of 94 ºC for 30 sec, 55 ºC for 30 sec, and 68 ºC for 60 sec (*see* **Note 10**); hold at 4 ºC.
3. Purification of the PCR amplicon by agarose gel electrophoresis (*see* **Note 2**) is recommended to remove *att*B primer-dimers.
4. To create the His_6-MBP fusion vector, the PCR product is recombined first into pDONR201 to yield an entry clone intermediate (BP reaction), and then into pDEST-HisMBP (LR reaction; *see* **Note 11**).
    a. Add to a microcentrifuge tube on ice: 300 ng of the PCR product in TE or H_2O, 300 ng of pDONR201 DNA, 4 μl of 5X BP reaction buffer, and enough TE to bring the total volume to 16 μl. Mix well.
    b. Thaw BP Clonase enzyme mix on ice (2 min) and then vortex briefly (2 sec) twice (*see* **Note 12**).
    c. Add 4 μl of BP Clonase enzyme mix to the components in (a.) and vortex briefly twice.
    d. Incubate the reaction at room temperature for at least 4 hours (*see* **Note 13**).

e. Add to the reaction: 1 μl of 0.75 M NaCl, 3 μl (ca. 450 ng) of the destination vector (pDEST-HisMBP), and 6 μl of LR Clonase enzyme mix (*see* **Note 12**). Mix by vortexing briefly.

f. Incubate the reaction at room temperature for 3-4 hrs.

g. Add 2.5 μl of the proteinase K stop solution and incubate for 10 min at 37°C.

h. Transform 2 μl of the reaction into 50 μl of chemically competent DH5α cells (*see* **Note 3**).

i. Pellet the cells by centrifugation, gently resuspend pellet in 100-200 μl of LB broth and spread on an LB agar plate containing ampicillin (100 μg/ml), the selective marker for pDEST-HisMBP (*see* **Fig. 1.**). Incubate the plate at 37 °C overnight (*see* **Note 14**).

5. Plasmid DNA is isolated from saturated cultures started from individual ampicillin-resistant colonies and screened by PCR, using the gene-specific primers N1 and C, to confirm that the clones have the expected structure. Alternatively, plasmids can be purified and screened by conventional restriction digests using appropriate enzymes. At this stage, we routinely sequence putative clones to ensure that there are no PCR-induced mutations.

3.2 Pilot Expression Experiment

Prior to large-scale cell growth and purification, the fusion protein is overproduced on a small scale to assess its solubility. The amount of fusion protein in the soluble fraction of the crude cell lysate is compared by SDS-PAGE with the total amount of fusion protein in the cells, and the results are analyzed by visual

inspection of the stained gel. In a parallel experiment, the fusion protein is cleaved *in vivo* to ascertain whether or not it is an efficient substrate for TEV protease and to evaluate the solubility of the passenger protein after it is cleaved from the His_6-MBP tag. If the passenger protein remains soluble after intracellular processing, then it is also likely to be soluble after the fusion protein has been purified and processed *in vitro*. Conversely, poor solubility after intracellular processing indicates that troubleshooting will be required before production can be scaled up (*see* **Subheading 3.2.6.**).

3.2.1. Selecting a Host Strain of E. coli

The production of TEV protease from the expression vector pRK603 (***4***) is initiated by adding anhydrotetracycline to the cell culture. This allows it to be regulated independently of the IPTG-inducible His_6-MBP fusion vector, which is important because sometimes delaying the induction of TEV protease until the fusion protein substrate has had time to accumulate in the cells results in greater solubility of the passenger protein after cleavage (***4,6***). To achieve regulated expression of TEV protease, the *in vivo* processing experiment must be performed in a strain of *E. coli* that produces the Tet repressor, such as BL21PRO or DH5PRO (Clontech, Palo Alto, CA, USA). We prefer BL21Pro because of its robust growth characteristics and the fact that it lacks two proteases (Lon and OmpT) that are present in most *E. coli* K12 strains.

3.2.2. Protein Expression

1. Transform competent BL21PRO or DH5PRO cells that already contain pRK603 (*see* **Notes 5** and **6**) with the His_6-MBP fusion protein expression vector and spread them on an LB agar plate containing ampicillin (100 g/ml) and kanamycin (25 g/ml). See **Subheadings 2.1.10.** and **2.2.3.** for the preparation of LB medium and antibiotic stock solutions. Incubate the plate overnight at 37 ºC.
2. Inoculate 2-5 ml of LB medium containing ampicillin (100 g/ml) and kanamycin (25 g/ml) in a culture tube or shake-flask with a single colony from the plate. Grow to saturation overnight at 37 ºC with shaking.
3. The next morning, inoculate 50 ml of the same medium in a 250 ml baffled-bottom flask with 0.5 ml of the saturated overnight culture.
4. Grow the cells at 37 ºC with shaking to mid-log phase (OD_{600nm} ~ 0.5).
5. Add IPTG (1 mM final concentration) and adjust the temperature to 30 ºC (*see* **Note 15**).
6. After 2 hours, divide the culture into two separate flasks (ca. 20 ml in each). Label one flask "+" and the other "–".
7. add anhydrotetracycline to the "+" flask (100 ng/ml final concentration).
8. After 2 more hours, measure the OD_{600nm} of the cultures (dilute cells 1:10 in LB to obtain an accurate reading). An OD_{600nm} of about 3-3.5 is normal, although lower densities are possible. If the density of either culture is much lower than this, it may be necessary to adjust the volume of the samples that are analyzed by SDS-PAGE.
9. Transfer 10 ml of each culture to a 15 ml conical centrifuge tube and pellet the cells by centrifugation at 4 °C.

10. Resuspend the cell pellets in 1 ml of lysis buffer (*see* **Subheading. 2.2.8.**) and then transfer the suspensions to a 1.5 ml microcentrifuge tube.
11. Store the cell suspensions at –80 ºC overnight. Alternatively, the cells can be disrupted immediately by sonication (after freezing and thawing) and the procedure continued without interruption, as described below.

3.2.3. Sonication and Sample Preparation

1. Thaw the cell suspensions at room temperature, then place them on ice.
2. Lyse the cells by sonication (*see* **Note 16**).
3. Prepare samples of the total intracellular protein from the "+" and "–" cultures (T+ and T–, respectively) for SDS-PAGE by mixing 50 µl of each sonicated cell suspension with 50 µl of 2x SDS-PAGE sample buffer containing 10% (v/v) 2-mercaptoethanol.
4. Pellet the insoluble cell debris (and proteins) by centrifuging the sonicated cell suspension from the each culture at maximum speed in a microcentrifuge for 10 min at 4 °C.
5. Prepare samples of the soluble intracellular protein from the "+" and "–" cultures (S+ and S–, respectively) for SDS-PAGE by mixing 50 µl of each supernatant from step 4 with 50 µl of 2x SDS-PAGE sample buffer containing 10% (v/v) 2-mercaptoethanol.

3.2.4. SDS-PAGE

We typically use precast Tris-Glycine or NuPAGE gradient gels for SDS-PAGE to assess the yield and solubility of MBP fusion proteins (*see* **Note 7**). Of course, the investigator is free to choose any appropriate SDS-PAGE formulation, depending on the protein size and laboratory preference.

1. Heat the T–, T+, S– and S+ protein samples at 90 ºC for about 5 min and then spin them at maximum speed in a microcentrifuge for 5 min.
2. Assemble the gel in the electrophoresis apparatus, fill it with SDS-PAGE running buffer, load the samples (10 μl each), and carry out the electrophoretic separation according to standard lab practices. T and S samples from each culture ("+" and "–") are loaded in adjacent lanes to allow easy assessment of solubility. Molecular weight standards may also be loaded on the gel, if desired.
3. Stain the proteins in the gel with GelCode® Blue reagent, PhastGel™ Blue R, or a suitable alternative.

3.2.5. Interpreting the Results

The MBP fusion protein should be readily identifiable in the T– sample after the gel is stained since it will normally be the most abundant protein in the cells. Molecular weight standards can also be used to corroborate the identity of the fusion protein band. If the S– sample contains a similar amount of the fusion protein, this indicates that it is highly soluble in *E. coli*. On the other hand, if little or no fusion protein is observed in the S– sample, then this is an indicator of poor solubility. Of course, a range of intermediate states is also possible.

If the fusion protein is an efficient substrate for TEV protease, then little of it will be present in the T+ and S+ samples. Instead, one should observe a prominent band at ca. 42 kDa that corresponds to the His_6-MBP moiety and another prominent band migrating with the expected mobility of the passenger protein. If the fusion protein is a poor substrate for the protease, then the "+" samples will look similar to the "–" samples.

If the passenger protein is soluble after it is released from His_6-MBP, then a similar amount will be present in the T+ and S+ lanes. On the other hand, some or all of the passenger protein may precipitate at this stage. If a substantial fraction of the passenger protein is insoluble, then troubleshooting may be necessary (*see* **Subheading 3.2.6.**). Alternatively, an acceptable yield might still be obtained by scaling up cell production.

An example of a pilot expression experiment is shown in **Fig. 3**. In this case the fusion protein (MBP-Lon) was highly soluble and readily cleaved *in vivo* by TEV protease. Note also that the Lon protease catalytic domain remained soluble after it was cleaved from the dual His_6-MBP tag.

3.2.6. Troubleshooting

Not every MBP fusion protein will be highly soluble. However, solubility usually can be increased by reducing the temperature of the culture from 30 °C to 25 °C or even lower during the time that the fusion protein is accumulating in the cells (*i.e.*, after the addition of IPTG). In some cases, the improvement can be quite dramatic. It may also be helpful to reduce the IPTG concentration to a level that will result in partial induction of the fusion protein. The appropriate IPTG concentration must be determined empirically,

but is generally in the range of 10-20 □M. Under these conditions, longer induction times (18-24 hrs) are required to obtain a reasonable yield of fusion protein.

If the fusion protein is a poor substrate for TEV protease *in vivo*, then the same is likely to be true *in vitro*. However, in most cases it is still possible to obtain a sufficient quantity of the pure passenger protein by scaling up production (e.g., from 4 to 6 liters of cells or more). In especially problematic cases, the efficiency of the protease digest can be improved by inserting additional amino acid residues between the TEV protease recognition site and the N-terminus of the passenger protein. We have used both polyglycine and a FLAG-tag epitope in this position with good results (***7,8***).

Occasionally, a passenger protein may accumulate in a soluble but biologically inactive form after intracellular processing of an MBP fusion protein. Exactly how and why this occurs is unclear, but we suspect that fusion to MBP somehow enables certain proteins to evolve into kinetically trapped, folding intermediates that are no longer susceptible to aggregation. Therefore, although solubility after intracellular processing is a useful indicator of a passenger protein's folding state in most cases, it is not absolutely trustworthy. For this reason, we strongly recommend that a biological assay be employed (if available) at an early stage to confirm that the passenger protein is in its native conformation.

3.3. Large Scale Cell Growth and Protein Purification

3.3.1. Cell Growth

1. Transform competent BL21 cells (Novagen, Madison, WI, USA) with the His_6-MBP-passenger expression vector and select transformants on LB agar plates

containing 100 μg/ml ampicillin (*see* **Subheading 2.1.10.** for recipes).

Alternatively, if the passenger protein ORF contains codons that are rarely used in *E. coli* (*see* **Note 6**) then tranform BL21 CodonPlus™ RIL (Stratagene, La Jolla, CA, USA) or Rosetta™ (Novagen) cells instead, and select transformants on LB agar plates containing both ampicillin (100 μg/ml) and chloramphenicol (30 μg/ml). Make a stock solution of 30 mg/ml chloramphenicol in ethanol and store at -20 ºC for up to 6 months.

2. Inoculate 100 ml of LB broth in a 500 ml baffle-bottom shake flask (Bellco Glass, Vineland, NJ, USA, or the equivalent) containing the appropriate antibiotic(s) with a single drug-resistant colony from the transformation in step 1. Shake overnight at 37 °C.
3. Add 10 ml of the saturated overnight culture to 1 liter of fresh LB broth plus antibiotic(s) in a 4 liter baffle-bottom shake flask. To ensure that there will be an adequate yield of pure protein at the end of the process, we ordinarily grow at least 4 liters of cells at a time. Sterile glucose can be added to 0.2% to increase biomass production. Make a stock solution of 20% (w/v) glucose in H_2O, filter sterilize, and store at 4 °C. Shake the flask(s) at 37 °C (ca. 250 rpm) until the cells reach mid-log phase (OD_{600nm} ~ 0.5).
4. Shift the temperature to 30 °C **(***see* **Note 17)** and then add IPTG to a final concentration of 1 mM. Continue shaking for 4-6 hrs.
5. Recover the cells by centrifugation. Freeze the cell pellet(s) at –80 °C.

3.3.2. Protein Purification

1. All procedures are performed at 4-8°C. Thaw the cell pellet(s) on ice and suspend in ice-cold 25 mM HEPES (pH 8), 200 mM sodium chloride, 25 mM imidazole buffer containing protease inhibitors, using at least 10 ml per gram wet weight of cell paste (*see* **Note 18**). Lyse the cells and clarify the crude cell extract by high-speed centrifugation and filtration (*see* **Note 19**).
2. Apply the supernatant to a column of Ni-NTA resin (Qiagen, Valencia, CA, USA) equilibrated in 25 mM HEPES (pH 8), 200 mM sodium chloride, 25 mM imidazole buffer without protease inhibitors (*see* **Note 18** and **20**). Wash the column with this buffer until a stable baseline is reached and then elute the bound fusion protein (His_6-MBP-passenger) with a linear gradient over ten column volumes into 25 mM HEPES (pH 8), 200 mM sodium chloride, 250 mM imidazole buffer (*see* **Note 18**). The fusion protein usually elutes between 100-150 mM imidazole. The purity of the fusion protein at this stage is typically in the range of 70-80% (*see* **Fig. 4**, lane 3).
3. To reduce the imidazole concentration to ca. 25 mM, the peak fractions containing the fusion protein are pooled and diluted with 25 mM HEPES (pH 8), 200 mM sodium chloride buffer that does not contain protease inhibitors or imidazole (*see* **Notes 18** and **21**). The concentration of the fusion protein is estimated by measuring the A_{280}.
4. His_6-TEV protease is added to the pooled, diluted fractions from step 3 and incubated overnight at 4°C (*see* **Note 22**). The progress of the cleavage reaction can be monitored by SDS-PAGE (*see* **Fig. 4**, lane 4).

5. The His_6-TEV protease-treated pool from step 4 is applied to a second Ni-NTA column under the same conditions used during application to the first column (*see* **Note 23**). Pure passenger protein passes through the Ni-NTA column while residual undigested fusion protein, His_6-MBP, His_6-TEV protease, and any contaminants are retained (*see* **Fig. 4**, lanes 5 and 6). The purity of the passenger protein at this stage is greater than 90% and is of sufficient quality for crystallization trials (*see* **Note 24** and **Fig. 5**).

4. Notes

1. We recommend a proofreading polymerase such as *Pfu* Turbo (Stratagene, La Jolla, CA, USA), Platinum *Pfx* (Invitrogen, Carlsbad, CA, USA), or Deep Vent (New England Biolabs, Beverly, MA, USA) to minimize the occurrence of mutations during PCR.
2. We typically purify fragments by horizontal electrophoresis in 1% agarose gels run in TAE buffer (40 mM Tris-acetate, pH 8.0, 1 mM EDTA). It is advisable to use agarose of the highest possible purity (*e.g.* Seakem-GTG from FMC BioPolymer, Philadelphia, PA, USA). Equipment for horizontal electrophoresis can be purchased from a wide variety of scientific supply companies. DNA fragments are extracted from slices of the ethidium bromide-stained gel using a QIAquick™ gel extraction kit (Qiagen, Valencia, CA, USA) in accordance with the instructions supplied with the product.
3. Any *gyrA*$^+$ strain of *E. coli* can be used. We prefer competent DH5 cells (Invitrogen, Carlsbad, CA, USA) because they are easy to use and very efficient.

4. We prefer the Wizard® miniprep kit (Promega, Madison, WI, USA) or the QIAprep™ Spin miniprep kit (Qiagen, Valencia, CA, USA), but similar kits can be obtained from a wide variety of vendors.

5. While any method for the preparation of competent cells can be used (*e.g.* $CaCl_2$), we prefer electroporation because of the high transformation efficiency that can be achieved. Detailed protocols for the preparation of electroccompetent cells and electrotransformation procedures can be obtained from the electroporator manufacturers (*e.g.* Bio-Rad, BTX, Eppendorf). Briefly, the cells are grown in 1 liter of LB medium (with antibiotics, if appropriate) to mid-log phase (OD_{600} ~0.5) and then chilled on ice. The cells are pelleted at 4 ºC, resuspended in 1 liter of ice-cold H_2O and then pelleted again. After several such washes with H_2O, the cells are resuspended in 3-4 ml of 10% glycerol, divided into 50 μl aliquots, and then immediately frozen in a dry ice/ethanol bath. The electrocompetent cells are stored at -80 ºC. Immediately prior to electrotransformation, the cells are thawed on ice and mixed with 10-100 ng of DNA (*e.g.* a plasmid vector or a Gateway® reaction). The mixture is placed into an ice-cold electroporation cuvette and electroporated according to the manufacturers recommendations (*e.g.* a 1.8 kV pulse in a cuvette with a 1 mm gap). One ml of SOC medium (***9***) is immediately added to the cells and they are allowed to grow at 37 ºC with shaking (ca. 250 rpm) for 1 hr. 5-200 μl of the cells are then spread on an LB agar plate containing the appropriate antibiotic(s).

6. If the open reading frame encoding the passenger protein contains codons that are rarely used in *E. coli*, this can adversely affect the yield of an MBP fusion protein.

In such cases, it is advisable to introduce an additional plasmid into the host cells that carries the cognate tRNA genes for rare codons. The pRIL plasmid (Stratagene, La Jolla, CA, USA) is a derivative of the p15A replicon that carries the *E. coli argU*, *ileY*, and *leuW* genes, which encode the cognate tRNAs for AGG/AGA, AUA, and CUA codons, respectively. pRIL is selected for by resistance to chloramphenicol. In addition to the tRNA genes for AGG/AGA, AUA, and CUA codons, the pRARE accessory plasmid in the Rosetta™ host strain (Novagen, Madison, WI, USA) also includes tRNAs for the rarely used CCC and GGA codons. Like pRIL, the pRARE plasmid is a chloramphenicol-resistant derivative of the p15A replicon. Both of these tRNA accessory plasmids are compatible with derivatives of pDEST-HisMBP. On the other hand, they are incompatible with the vector pRK603 that we use for intracellular processing experiments (*see* **Subheading 3.2**). Nevetheless, because pRK603 and the tRNA accessory plasmids have different antibiotic resistance markers, it is possible to force cells to maintain both plasmids by simultaneously selecting for kanamycin and chloramphenicol resistance. Alternatively, the kanamycin-resistant TEV protease expression vector pKM586, a pRK603 derivative with the replication machinery of a pSC101 replicon, can be stably maintained in conjunction with p15A-type tRNA plasmids.

7. We find it convenient to use precast gels for SDS-PAGE gels (*e.g.* 1.0 mm x 10 well, 10-20% Tris-glycine gradient), running buffer, and electrophoresis supplies from Invitrogen (Carlsbad, CA, USA).

8. Alternatively, the PCR reaction can be performed in two separate steps, using primers N1 and C in the first step and primers N2 and C in the second step. The PCR amplicon from the first step is used as the template for the second PCR. All primers are used at the typical concentrations for PCR in the two-step protocol.
9. The PCR reaction can be modified in numerous ways to optimize results, depending on the nature of the template and primers. See (***9***) (Vol. 2, Chapter 8) for more information.
10. PCR cycle conditions can also be varied. For example, the extension time should be increased for especially long genes. A typical rule-of-thumb is to extend for sixty seconds per kilobase of DNA.
11. This "one-tube" Gateway® protocol bypasses the isolation of an "entry clone" intermediate. However, the entry clone may be useful if the investigator intends to experiment with additional Gateway® destination vectors, in which case the BP and LR reactions can be performed sequentially in separate steps; detailed instructions are included with the Gateway® PCR kit. Alternatively, entry clones can easily be regenerated from expression clones via the BP reaction, as described in the instruction manual.
12. Clonase enzyme mixes should be thawed quickly on ice and then returned to the -80 ºC freezer as soon as possible. It is advisable to prepare multiple aliquots of the enzyme mixes the first time that they are thawed in order to avoid repeated freeze-thaw cycles.
13. At this point, we remove a 5 □l aliquot from the reaction and add it to 0.5 □l of proteinase K stop solution. After 10 min at 37 °C, we transform 2 □l into 50 □l of

competent DH5 cells (*see* **Note 3**) and spread 100-200 l on an LB agar plate containing kanamycin (25 g/ml), the selective marker for pDONR201. From the number of colonies obtained, it is possible to gauge the success of the BP reaction. Additionally, entry clones can be recovered from these colonies in the event that no transformants are obtained after the subsequent LR reaction.

14. If very few or no ampicillin-resistant transformants are obtained after the LR reaction, the efficiency of the process can be improved by incubating the BP reaction overnight.
15. 30 ºC is the optimum temperature for TEV protease activity. At 37 ºC, the protease does not fold properly in *E. coli* and little processing will occur. Reducing the temperature also improves the solubility of some MBP fusion proteins.
16. We routinely break cells in a 1.5 ml microcentrifuge tube on ice with two or three 30 sec pulses using a VCX600 sonicator (Sonics & Materials, Newtown, CT, USA) with a microtip at 38% power. The cells are cooled on ice between pulses.
17. We have found that decreasing the induction temperature to 30 °C increases the quality and solubility of the fusion protein without significantly decreasing the yield, especially in the presence of glucose.
18. Buffers compatible with IMAC [HEPES, phosphate and Tris buffers (25-50 mM, pH 7-8)] must be used. We include sodium chloride (100-500 mM) and a low concentration of imidazole (25 mM) in our cell lysis and IMAC equilibration buffers to decrease non-specific adsorption to the resin. In addition, protease inhibitors [PMSF or AEBSF (1 mM) or Complete EDTA-free protease inhibitor cocktail tablets and benzamidine hydrochloride (1 mM)] are added to the cell lysis

buffer unless they are contraindicated (e.g., for the purification of an **active** recombinant protease, the inhibitors may have to be excluded). We always avoid EDTA and other divalent metal chelators in our cell lysis and IMAC buffers. The 25 mM HEPES (pH 8), 200 mM sodium chloride, 25 mM imidazole buffer used for cell lysis and IMAC is prepared by mixing 5.96 g of HEPES, 11.69 g of sodium chloride and 1.70 g of imidazole with distilled water to ca. 950 ml. The solution is adjusted to pH 8 with 10 *N* sodium hydroxide and the volume increased to 1 liter with H_2O. The buffer is filtered through a 0.22 μm polyethersulfone membrane and stored at 4ºC. Protease inhibitors (for cell lysis buffer only) are added immediately before use (7.83 mg benzamidine hydrochloride and 1 tablet of Complete EDTA-free protease inhibitor cocktail per 50 ml of buffer). The 25 mM HEPES (pH 8), 200 mM sodium chloride buffer is prepared in the same manner, except that the imidazole is omitted. The 25 mM HEPES (pH 8), 200 mM sodium chloride, 250 mM imidazole buffer used for IMAC is prepared by mixing 5.96 g of HEPES, 11.69 g of sodium chloride and 17 g of imidazole with distilled water to ca. 950 ml. The solution is adjusted to pH 8 with hydrochloric acid and then the volume is increased to 1 liter with H_2O. The buffer is filtered through a 0.22 μm polyethersulfone membrane and stored at 4ºC.

19. For large-scale protein purification we routinely break cells using a Gaulin APV fluidizer at 10-11,000 psi for 2-3 rounds. Other homogenization techniques such as French press, sonication, or manual shearing should yield comparable results. Centrifugation of the disrupted cell suspension for at least 30 min at 30,000 x *g* is recommended. Filtration through a 0.2 μm polyethersulfone or cellulose acetate

membrane is helpful to remove residual particulates and fines prior to chromatography.

20. We use Ni-NTA Superflow (Qiagen, Valencia, CA, USA) and an Amersham Biosciences ÄKTA Explorer chromatography system. Ten to 25 ml columns are employed, depending on the amount of fusion protein produced by the cells. A properly poured 10 ml Ni-NTA Superflow column (in an Amersham Biosciences column XK16/20) can be run at 2-4 ml/min (backpressure less than 0.4 MPa) and will bind up to 100 mg of fusion protein. If a chromatography system is not available, the IMAC can be performed using a peristaltic pump or manually by gravity. If the latter is used, Ni-NTA agarose should be substituted for Superflow and the elution performed with step increases of imidazole in 25 mM increments. Binding and elution profiles can be monitored spectrophotometrically at 280 nm and by SDS-PAGE. Care must be taken to properly zero the spectrophotometer because imidazole has significant absorption in the UV range.

21. It is convenient to reduce the imidazole concentration to 25 mM at this step in preparation for the second round of IMAC. If the volume is too large, it can be reduced using any commercially available concentrating units (e.g., an Amicon Stir-Cell Concentrator with a YM membrane). Reduction of volume is not critical, however, because IMAC and the subsequent cleavage reaction (*see* **Note 22**) are insensitive to this variable.

22. His_6-TEV protease is active in all buffers compatible with IMAC. We use 1 mg of protease per ca. 150 mg of fusion protein (estimated by measuring the A_{280}). Using more His_6-TEV protease than required will have no deleterious effect. Volumes as

large as 500 ml can be used. Incubations can be performed between 4-30°C with equivalent results, although an overnight incubation at 4°C is most convenient. His_6-TEV protease can be purchased from Invitrogen (Carlsbad, CA, USA). However, for large-scale applications it is far more cost effective to overproduce and purify the enzyme in-house as described (***5***).

23. An Ni-NTA column of equal or greater volume relative to the first column is required. After dilution of the imidazole to 25 mM, all other buffer components should be the same. In instances where saturation of the Ni-NTA resin during the first round of IMAC occurs, then a column of greater volume is recommended to avoid contamination of the passenger protein emerging in the column effluent during the second round of IMAC.
24. While the passenger protein is typically free of contaminants after the second IMAC step, it may still exist as a mixture of oligomeric forms including some high molecular weight aggregates. For this reason, we recommend that size exclusion (gel filtration) chromatography be employed as a polishing step. This is also a good way to exchange the protein into an appropriate buffer for crystallizaion trials. Additionally, we recommend that the molecular weight of the final product be verified by electrospray mass spectrometry if possible.

Acknowledgements

We wish to express our gratitude to Dr. Istvan Butros for providing us with a photograph of Lon protease crystals (*see* **Fig. 5**). We also wish to thank Rachel Kapust, Karen Routzahn, and Jeff Fox for their valuable contributions to the development of these

methods and Invitrogen, Inc. for granting us early access to Gateway® Cloning technology.

References

1. Kapust, R. B. and Waugh, D. S. (1999) *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci.* **8,** 1668-1674.

2. Fox, J. D., Routzahn, K. M., Bucher, M. H. and Waugh, D. S. (2003) Maltodextrin-binding proteins from diverse bacteria and archaea are potent solubility enhancers. *FEBS Letts.* **537**, 53-57.

3. Routzahn, K. M. and Waugh, D. S. (2002) Differential effects of supplementary affinity tags on the solubility of MBP fusion proteins. *J. Struct. Funct. Genomics* **2**, 83-92.

4. Kapust, R. B. and Waugh, D. S. (2000) Controlled intracellular processing of fusion proteins by TEV protease. *Protein Expr. Purif.* **19**, 312-318.

5. Kapust, R. B., Tozser, J., Fox, J. D., Anderson, D. E., Cherry, S., Copeland, T. and Waugh, D. S. (2001). Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic efficiency. *Protein Eng.* **14**, 993-1000.

6. Fox, J. D. and Waugh, D. S. (2003). Maltose-binding protein as a solubility enhancer. *Methods Mol. Biol.* **205**, 99-117.

7. Evdokimov, A. G., Tropea, J. E. Routzahn, K. M., Copeland, T. D. and Waugh, D. S. (2001). Structure of the N-terminal domain of *Yersinia pestis* YopH at 2.0 Å resolution. *Acta. Crystallogr. D Biol. Crystallogr*. **57**, 793-799.

8. Evdokimov, A. G., Tropea, J. E., Routzahn, K. M. and Waugh, D. S. (2002). Three-dimensional structure of the type III secretion chaperone SycE from *Yersinia pestis*. *Acta Crystallogr. D Biol. Crystallogr.* **58**, 398-406.

9. Sambrook, J. and Russell, D. W. (2001) *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Figure Legends

Fig. 1. Schematic representation of the Gateway® destination vector pDEST-HisMBP. This vector can be recombined with an entry vector that contains an ORF of interest, via the LR reaction, to generate a His_6-MBP fusion protein expression vector.

Fig. 2. Construction of a His_6-MBP fusion vector using PCR and Gateway® cloning technology. The ORF of interest is amplified from the template DNA by PCR, using primers N1, N2, and C. Primers N1 and C are designed to base-pair to the 5' and 3' ends of the coding region, respectively, and contain unpaired 5' extensions as shown. Primer N2 base-pairs with the sequence that is complementary to the unpaired extension of primer N1. The final PCR product is recombined with the pDONR201 vector to generate an entry clone, via the BP reaction. This entry clone is subsequently recombined with pDEST-HisMBP using LR Clonase to yield the final His_6-MBP fusion vector.

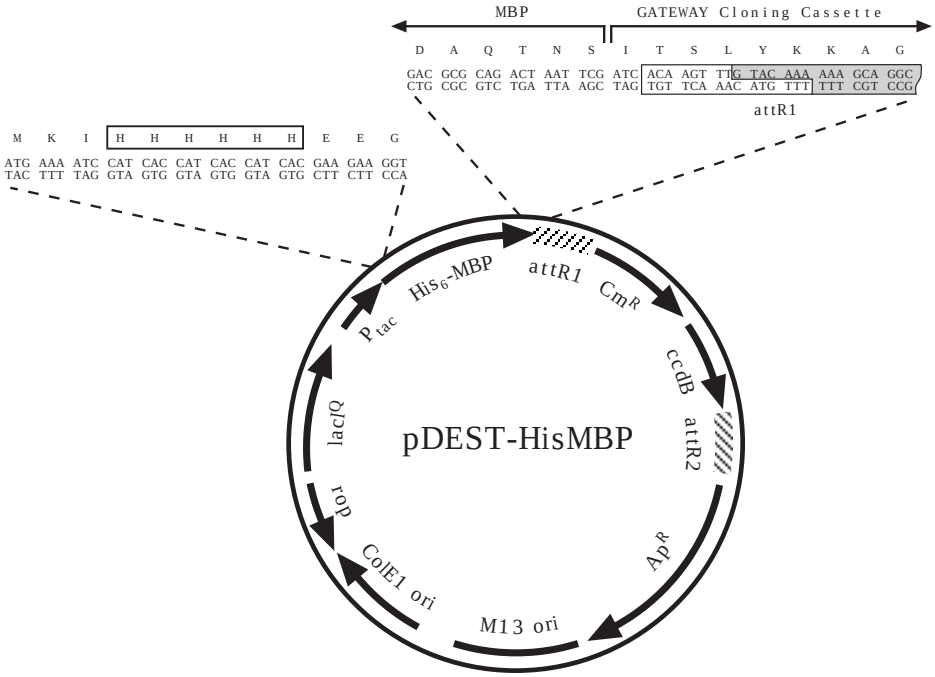
Fig. 3. Intracellular processing of a His_6-MBP fusion protein by TEV protease. The catalytic domain of *E. coli* Lon protease was expressed from a derivative of pDEST-HisMBP in BL21PRO cells that also contained the TEV protease expression vector pRK603 as described (*see* **Subheading 3.2.**). Lane1: molecular weight standards. Lane2: T–. Lane 3: S–. Lane 4: T+. Lane 5, S+ (*see* **Subheadings 3.2.4.** and **3.2.5.**).

Fig. 4. Generic purification of *E. coli* Lon protease catalytic domain (*see* **Subheading 3.3.**). Lane 1: molecular weight standards. Lane 2: Soluble cell extract. Lane 3: Peak

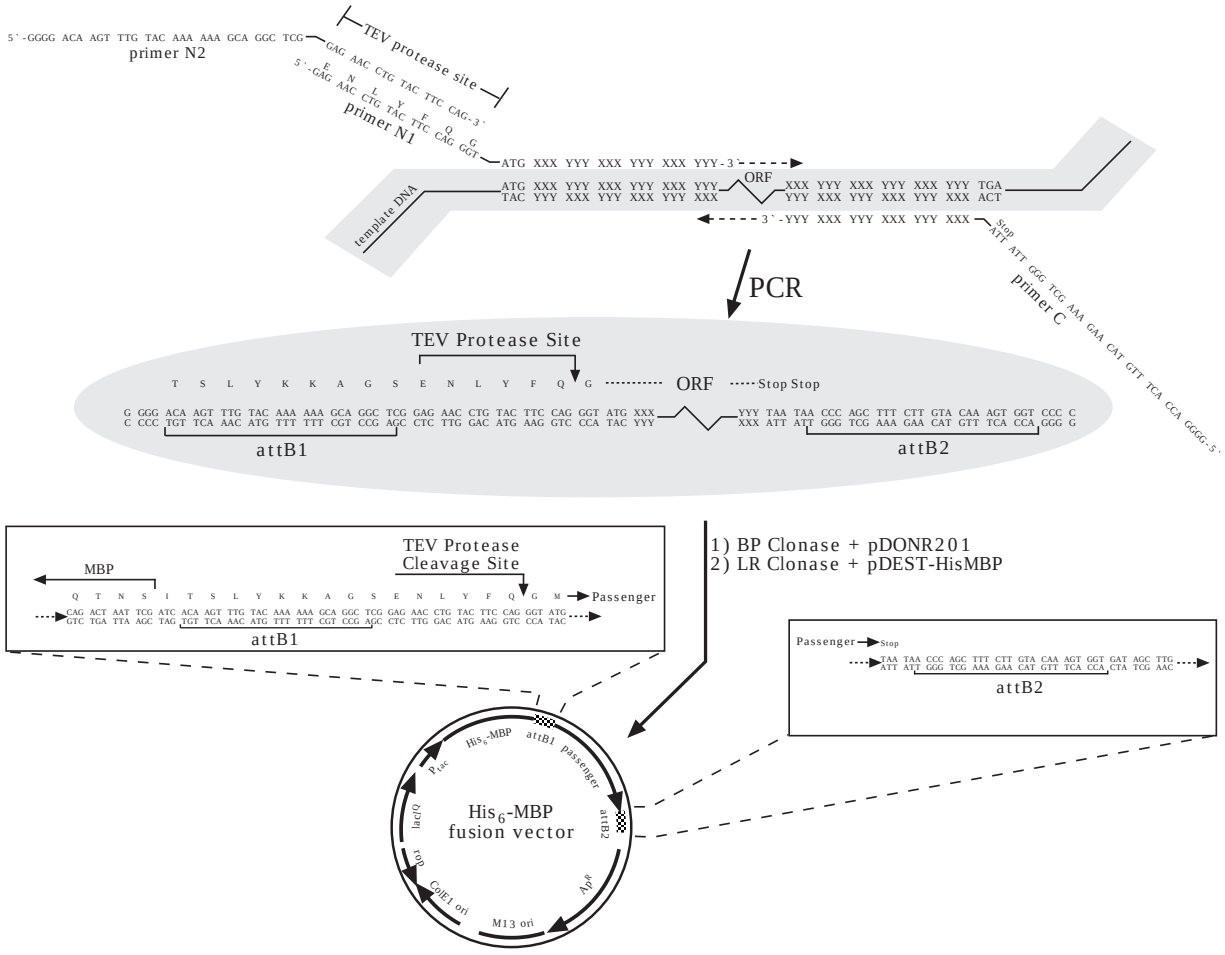
fractions from the first round of IMAC, eluted with an imidazole gradient. Lane 4: TEV protease digest. Lane 5: flow-through fractions after second round of IMAC. Lane 6: material absorbed to Ni-NTA column during second round of IMAC.

Fig. 5. Crystals of *E. coli* Lon protease catalytic domain obtained after the protein was purified as described (*see* **Subheading 3.3.** and **Note 24**). Space group: $P3_2$. Diffraction limit: 1.75 Å (beam line X-9B at NSLS, Brookhaven, NY).

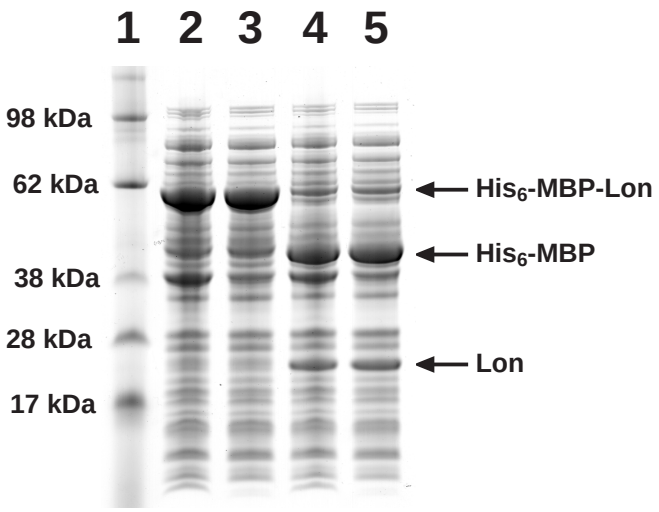
Tropea *et al.*, Fig. 1



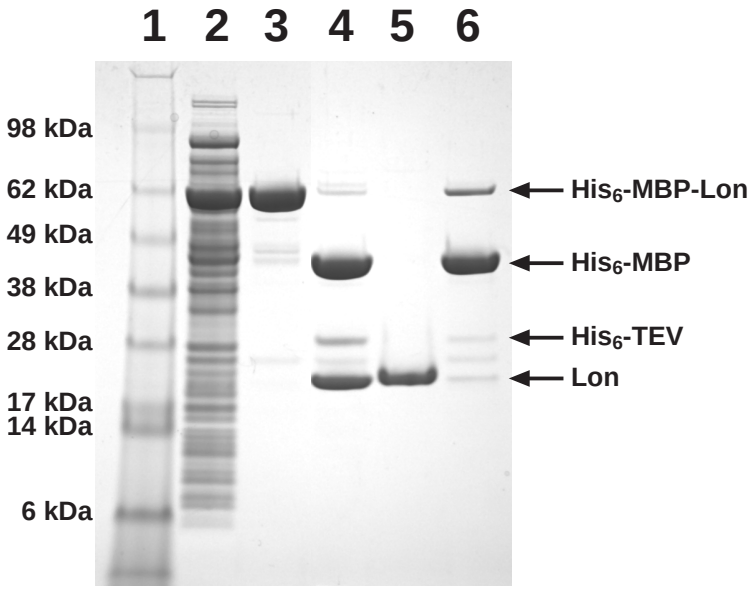
Tropea *et al.*, Fig. 2



Tropea *et al.*, Fig. 3



Tropea *et al.*, Fig. 4



Tropea *et al.*, Fig. 5